# A Multidimensional Histogram Rain-Flagging Technique For SeaWinds on QuikSCAT

**James N. Huddleston and Bryan W. Stiles**
**Jet Propulsion Laboratory, MS 300-319**
**California Institute of Technology**
**4800 Oak Grove Drive, Pasadena, CA 91109, USA**
**Tel: 818.354.1748, Fax: 818.393.5184**
**Email: James.N.Huddleston@jpl.nasa.gov**

*Abstract* — The SeaWinds scatterometer was developed by NASA JPL to measure the speed and direction of global ocean surface winds. SeaWinds was launched aboard the QuikSCAT spacecraft on June 19, 1999 and has continued to operate successfully since being turned on. Although the initial SeaWinds wind vector products were of excellent quality, they were occasionally degraded by the presence of rain. It soon became obvious that a way to flag wind vector cells for rain contamination was needed.

We have determined a set of parameters that are sensitive to rain and are computed from the scatterometer measurements. These parameters are: 1) the retrieved wind speed, 2) the retrieved wind direction relative to the satellite ground track, 3) the normalized beam difference, which indicates a statistically significant imbalance in the beams relative to the geophysical model function, 4) the maximum likelihood estimate calculated by the wind retrieval algorithm, and 5) the radiometric brightness temperature. Using these parameters and external rain information (SSM/I) an estimate of the conditional probability of rain given the parameters is developed using a multidimensional histogram technique. This probability estimate is then used to flag rain contaminated wind vector cells using only scatterometer data. This technique is currently employed to generate a rain flag for SeaWinds on QuikSCAT data.

In this paper, the effects of rain on SeaWinds data are explored, the rain flagging technique is explained, and the performance of the rain flag is illustrated using a number of metrics.

## INTRODUCTION

The SeaWinds on QuikSCAT scatterometer (QSCAT) was launched into earth orbit on June 19, 1999. Its mission: to determine the speed and direction of global ocean surface winds on a daily basis. It accomplishes this task by making multiple Ku-band measurements of the normalized radar cross section ($\sigma_0$) of the ocean's surface and using those measurements along with an empirical geophysical model function to infer the underlying wind. [1]

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

A quick perusal of the wind data products revealed regions of unexpectedly high wind speed. Rain was determined to be the culprit and a technique was needed for identifying and flagging wind vector cells that were likely to be contaminated by rain. Collocated SSM/I measurements were considered, but had two significant drawbacks. Firstly, they required the acquisition and ingestion of external data. Secondly, it was not possible to obtain collocated SSM/I measurements over the entire QSCAT swath without increasing the collocation time window beyond a reasonable value. For these reasons, it was very desirable to develop a rain-flagging technique that relied solely on QSCAT data.

## RAIN-SENSITIVE PARAMETERS

Our first task was to identify a set of parameters that were sensitive to rain and could be calculated using only QSCAT data. One of the most obvious parameters was the retrieved wind speed. When rain is present, a portion of the transmitted energy is reflected back at the instrument by the rain, thus causing an increase in the amount of energy measured by the scatterometer. This, in turn, produces retrieved winds having high speeds.

The scattering phenomenon of rain is believed to be isotropic. For this reason, the forward looking $\sigma_0$ measurements and the aft looking $\sigma_0$ measurements become nearly equal. According to the geophysical model function, this is consistent with wind oriented in a cross-swath direction. In other words, the presence of rain causes the retrieved wind to align with the cross-swath direction. We use the direction of the first ranked wind solution relative to the along track direction as our second rain sensitive parameter.

Another feature of rain, is that it causes $\sigma_0$ measurements to become inconsistent with the model function describing wind driven ocean. For example, we have noticed that rain will have a larger scattering effect on the measurements made by the H polarization inner beam than on measurements made by the V polarization outer beam. This causes an imbalance in the beams relative to the model function. If we compare the individual $\sigma_0$ values that we would expect to get, based on our retrieved wind, to those that we actually measure, we find that the inner beam measurements are higher than we expect and the outer beam measurements are lower than we expect. We have developed a third rain-sensitive parameter called the normalized beam difference (NBD) that quantifies this phenomenon in a statistically meaningful way:

$$NBD = \frac{\frac{1}{N}\sum_{i=1}^{N}\frac{\sigma_{0,i,meas}-\sigma_{0,i,model}}{\sigma_i}-\frac{1}{M}\sum_{j=1}^{M}\frac{\sigma_{0,j,meas}-\sigma_{0,j,model}}{\sigma_j}}{\sqrt{\frac{1}{N}+\frac{1}{M}}} \quad (1)$$

where N is the number of measurements on the inner beam, M is the number of measurements on the outer beam, $\sigma_{0,i,meas}$ is the i$^{th}$ measured $\sigma_0$ for the inner beam, $\sigma_{0,i,model}$ is the model $\sigma_0$ corresponding to the i$^{th}$ measured $\sigma_0$ for the inner beam, $\sigma_{0,j,meas}$ is the j$^{th}$ measured $\sigma_0$ for the outer beam, $\sigma_{0,j,model}$ is the model $\sigma_0$ corresponding to the j$^{th}$ measured $\sigma_0$ for the outer beam, $\sigma_i$ is the expected standard deviation of the i$^{th}$ measured $\sigma_0$ for the inner beam, and $\sigma_j$ is the expected standard deviation of the j$^{th}$ measured $\sigma_0$ for the outer beam. Model $\sigma_0$ values are determined by using the retrieved wind vector, the measurement geometry, and the geophysical model function. Standard deviations are estimated based on our measurement noise model [2].

Rain also causes $\sigma_0$ measurements to exhibit different statistics than those made of the ocean's surface under rain-free conditions. Rain-contaminated $\sigma_0$'s tend to have a higher variance. This phenomenon becomes evident when one examines the maximum likelihood objective function value calculated as part of the wind retrieval algorithm. When normalized by the number of $\sigma_0$ measurements, this fourth rain-sensitive parameter indicates the goodness of fit of the model function to the $\sigma_0$ measurements. Rain tends to produce smaller likelihood values.

Finally, radiometric brightness temperature have long been known to indicate rain and thus is our fifth rain-sensitive parameter. Through the efforts of Jones and Zek, the QSCAT data is now being processed into brightness temperatures [3].

Fig. 1 shows the NBD density for both rain-contaminated and rain-free wind vector cells. Both the rain-contaminated and rain-free curves have been normalized to have equal areas. In this figure, and for this entire study, rain was defined as an SSM/I derived integrated rain rate of > 2.0 km·mm/hr. Rain-free was defined as an SSM/I derived integrated rain rate of 0.0 km·mm/hr. In all cases, collocations with SSM/I were required to be within 30 minutes.[1] For parameters that rely on the retrieved wind vector, we have chosen to use the first ranked (most likely) of the multiple solution vectors that are identified by the wind retrieval algorithm. Fig. 2 shows a similar density plot for the swath-relative direction parameter. Important to note in these plots is that the rain-sensitive parameters are affected by the presence of rain. Rain tends to produce larger values of NBD and to produce swath-relative directions near 90°.

[1] The SSM/I data was kindly provided by Remote Sensing Systems and collocated by. Freilich and Vanhoff.

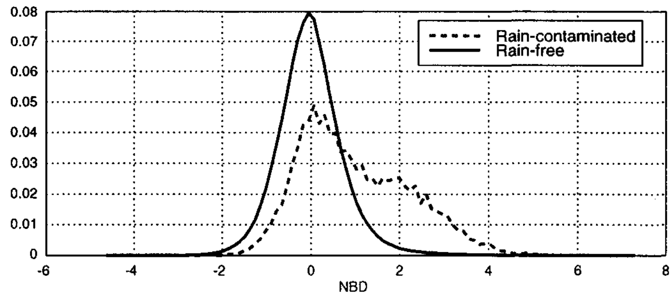


**Fig. 1.** Density of Normalized Beam Difference (NBD) for rain-contaminated and rain-free wind vector cells

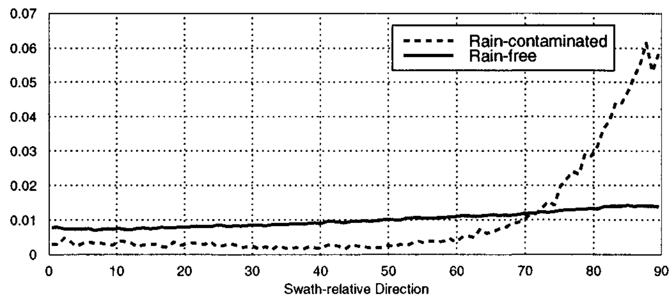


**Fig. 2.** Density of swath-relative direction for rain-contaminated and rain-free wind vector cells

## ALGORITHM OVERVIEW

Given our rain-sensitive parameters and a definition of rain based on the SSM/I integrated rain rate, it is straightforward to estimate the probability of rain as a function of those parameters. First, we create two multidimensional histograms, mapping a rain-sensitive parameter into each dimension of the histograms. In one of the histograms, we count the total number of wind vector cells. In the other histogram, we count the number of rain-contaminated wind vector cells. Simple division then yields what we are after: an estimate of the probability of rain versus the rain sensitive parameters.

It was not feasible to use all of the rain-sensitive parameters as histogram dimensions. For one, not all of the parameters are available over the entire QSCAT swath. As an example, the NBD parameter requires measurements from both beams in order to be calculated. Thus, it is not available in the outer swath where only measurements from the outer beam exist. The second reason for not mapping each rain-sensitive parameter into a dimension of the histograms is that as the dimensionality of the histograms get larger, it requires significantly more data to estimate the probability of rain over the entire histogram space. We have restricted the histograms to four dimensions and estimate the probabilities for the two-beam case independently from the single-beam case. For the two-beam case, we use speed, MLE, NBD, and the H polarization brightness temperature. For the single-beam case, we use speed, swath-relative direction, MLE, and the V polarization brightness temperature.

## TRAINING DETAILS

To estimate the probability of rain versus the various rain-sensitive parameters, we trained using 487 orbits of QSCAT data and SSM/I data collocated within 30 minutes. We required there to be 50 wind vector cells in a histogram bin in order to calculate a probability of rain. We applied an adaptive resolution technique so that if a bin contained less than 50 wind vector cells, the wind vector cells from nearby bins would be combined to form a probability estimate for the underfilled bin. A range of thresholds in probability were selected to causes various percentages of data to be flagged as rain-contaminated.

## PERFORMANCE ESTIMATES

We applied the MUDH rain flagging technique to 417 orbits of data which did not overlap with the training set. Several metrics were developed to evaluate the performance of the MUDH rain flag. The first is the "missed rain percentage" which is defined as the percentage of SSM/I data with an integrated rain rate > 2.0 km·mm/hr that was claimed to be rain-free by the MUDH algorithm. The second metric is the "false alarm percentage" which is defined as the percentage of SSM/I data with an integrated rain rate of 0.0 km·mm/hr that was claimed to be rain-contaminated by the MUDH algorithm.

Fig. 3 shows the false alarm and missed rain percentage of the MUDH rain flag versus the percent flagged as rain-contaminated. It is not surprising that as the algorithm is permitted to flag more data as rain-contaminated, it is less likely to miss rain and also more likely to falsely flag rain-free data as rain-contaminated.

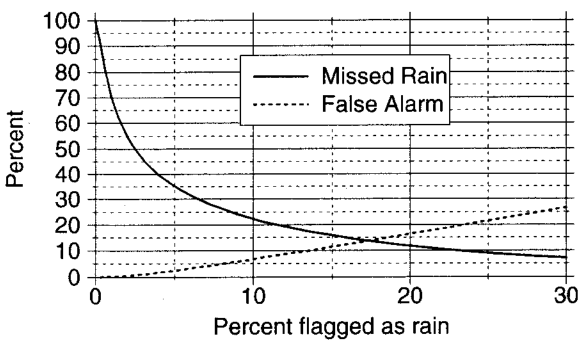


**Fig. 3.** MUDH rain flag performance versus percent flagged as rain

Wind retrieval metrics were also calculated for both the rain-contaminated and the rain-free data. We used ECMWF winds as "truth" and calculated the RMS speed and RMS direction errors for various ECMWF wind speed ranges. Table 1 indicates the differences between the rain-contaminated wind errors and the rain-free wind errors. For example, when examining ECMWF wind speed ranges from 3 to 7 m/s, the data flagged as rain-contaminated by the MUDH algorithm has a 4.88 m/s larger RMS speed error than does the data flagged as rain-free.

**Table 1.** Performance difference between rain-contaminated data and rain-free data

| ECMWF wind speed range (m/s) | RMS speed error difference (m/s) | RMS direction error difference (degrees) |
|---|---|---|
| 3 – 7 | 4.88 | 24.4 |
| 7 – 15 | 2.41 | 13.5 |
| 15 – 30 | 0.32 | 4.8 |

## CONCLUSIONS

The presence of rain contaminates $\sigma_0$ measurements made by Ku-band scatterometers operating at high incidence angles. The MUDH rain flagging technique provides a method for flagging contaminated wind vector cells using only data from the scatterometer itself; no external data source is needed to apply the technique. The MUDH technique is shown, via metrics, to accurately detect and flag rain-contaminated wind vector cells. Cells that are flagged for rain tend to have significantly degraded performance (as compared to a model field such as ECMWF) indicating the utility of flagging rain-contaminated data so it can be eliminated from wind analyses.

We would also like to note that the technique itself is quite flexible and can easily incorporate newly developed rain-sensitive parameters.

## REFERENCES


[1] C. Wu, J. Graf, M. Freilich, D. G. Long, M. Spencer, W. Tsai, D. Lisman, and C. Winn, "The SeaWinds scatterometer instrument," in *Proc. Int. Geoscience and Remote Sensing Symp.*, Pasadena, CA, Aug. 8-12, 1994, pp. 1511-1515.

[2] W. Tsai, S. V. Nghiem, J. N. Huddleston, M. W. Spencer, B. W. Stiles, and R. D. West, "Polarimetric Scatterometry: A Promising Technique for Improving Ocean Surface Wind Measurements from Space," accepted for publication in *IEEE Tran. on Geoscience and Remote Sensing*, 2000.

[3] L. Jones, R. Merhershahi, J. Zec, D. Long, "QuikSCAT Radiometric Measurements and Calibration," *IGARSS 2000*